DOCUMENT RESUME

ED 047 050 UD 011 215

AUTHOR Oakland, Thomas; Weilert, Evelyn
TITLE The Effects of Test-Wiseness Materials on Standardized Test Performance of Preschool Disadvantaged Children.
PUB DATE Feb 71
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, N.Y., February 1971

EDRS PRICE EDRS Price MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS Culturally Disadvantaged, *Curriculum Development, *Curriculum Evaluation, Preschool Children, *Standardized Tests, *Statistical Analysis, Student Testing, Task Performance, Testing, *Testing Problems

ABSTRACT

This study was an attempt to examine certain variables which apply to test-wiseness factors in assessing preschool and primary grade children. The study identified specific abilities which appear to be prerequisite for taking standardized tests, and to design curricular materials to facilitate the development of these abilities. It was found that the curricular test-wiseness materials had limited value in facilitating performance on the standardized test. However, teachers reported that the majority of children in the experimental group appeared to be more confident on paper and pencil tasks, and to remain task-oriented for longer periods of time. [Not available in hard copy due to marginal legibility of original document.] (Author/DM)

ED047050

The Effects of Test-Wiseness Materials on Standardized Test Performance of Preschool Disadvantaged Children

Thomas Oakland and Evelyn Weilert
The University of Texas at Austin

UD011215

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

This paper was presented at the annual convention of the American Educational Research Association in New York City, February, 1971.

Various criticisms have been directed toward the use of standardized tests with disadvantaged children (Anastasi, 1958; Deutsch and Fishman, 1964; Eels, et al., 1951; Haggard, 1954; Masland, et al., 1958; Tyler, 1956). It is generally agreed that the use of tests for classification and identification is appropriate for middle-class children on whom the norms are based, but these same tests may not equitably classify those lower-class and ethnically different children of average ability who simply are unfamiliar with the format and language of standardized tests. Some school systems have chosen to reject categorically the use of standardized tests with disadvantaged children because they tend to discriminate along socioeconomic and ethnic variables. While this practice may have some merit, categorical rejections of standardized tests because they discriminate is unjustified in that the value of norm-referenced tests lies in their ability to discriminate the extent to which people differentially manifest some ability or trait. The basic issue is the extent to which a test validly discriminates on only those variables it was designed to measure.

Newland (1963) has proposed that a set of basic assumptions underlying assessment must be met if test results are to be used with confidence. Implicit in these is the assumption that children possess certain abilities which are prerequisite to taking (group) standardized tests. Standardized tests used with primary-grade

children often require them to know basic concepts and to follow specific oral directions in order to take the test successfully. For example, it is not uncommon for readiness tests to use concepts such as left, right, row, column, most like, and opposite in attempting to assess perceptual, cognitive, and psychomotor abilities. The assessment of these latter abilities will be clouded if children lack an adequate understanding of these concepts and other prerequisite abilities.

Familiarity with the format and language used in standardized tests and possession of the abilities which are prerequisite to taking standardized tests pertain to a person's test wiseness. Test wiseness is defined as the ability to manifest test-taking skills which utilize the characteristics and format of a test and/or test-taking situation in order to receive a score commensurate with the abilities being measured.* Deficiencies in one or more of the abilities prerequisite to taking a test will attenuate the results, thereby depreciating the test's effectiveness to discriminate validly only those variables it was designed to measure.

Ebel (1965, p. 206) has stated that "...more error in measurement is likely to originate from students who have too little, rather than too much, skill in taking tests." Therefore, efforts directed toward increasing students' skills in taking tests should result in decreasing the error score and improving the precision of test

---

*This definition differs from the more commonly used definition of test wiseness: a subject's capacity to utilize the characteristics and formats of the test and/or the test taking situation to receive a high score (Millman, Bishop, and Ebel, 1965, p. 707).

interpretation. Previous investigations have attempted to facilitate the development of test-wiseness skills (Moore, Schultz, and Baker, 1966; Moore, 1968; Walstrom and Boersma, 1968; Slakter and Koehler, 1969) or to identify the variables associated with test wiseness. For example, Kreit (1967) examined the effects of test practice on the acquisition of test-taking skills of third-grade pupils. Four different group intelligence tests were administered to the experimental group, while the control group received only pre- and post-tests. Significant improvement in test-taking skills was apparent for the experimental group, presumably due to its increased exposure to a variety of tests. The relationship between intelligence and increased test wiseness was not significant. Millman and Setijari's (1966) study demonstrated the disadvantages under which students operate when taking a test with an unfamiliar format. Comparisons were made of the performance of Indonesian and American students on tests using open-ended and multiple choice questions involving arithmetic computation and vocabulary. The Indonesian students did relatively less well on the multiple choice items than they did on the open-ended questions. This differential performance presumably was due to their having had no prior experience with multiple choice items. The authors believe that the superior test wiseness of American students resulted in their having a significant advantage over the Indonesian students on the multiple choice questions. Thus, test wiseness may be an important source of variance when comparing the scores earned by two groups whose test-taking experience differs markedly.

All previous studies on test wiseness have dealt with Ss in the upper elementary grades or above. The present study was an attempt to examine certain variables which apply to test wiseness factors in assessing preschool and primary-grade children. The initial phase of the study was designed to identify specific abilities which appear to be prerequisite for taking standardized readiness tests. Having identified these abilities, curricular materials were designed to facilitate the development of these abilities. Evaluation of these curricular materials comprises the major focus of this article.

Method

Various readiness tests (Metropolitan Readiness Tests, Harrison-Stroud Reading Readiness Profiles, Lee-Clark Reading Readiness Test, Monroe Reading Aptitude Tests, and Murphy-Durrell Diagnostic Reading Readiness Test) were reviewed in an attempt to determine the abilities which appear to be prerequisite for taking readiness tests. The results of this review, together with teachers' comments, formed the basis for designing curricular materials that would increase the test wiseness of children unfamiliar with standardized tests and develop the abilities which appear to be prerequisite for taking readiness tests. The materials do the following:

(1) Begin with few items and options per page and gradually increase them in number until the page is similar in appearance to an actual test page.

(2) Provide practice in working in columns and rows.

(3) Teach the concepts right, left, up, down, opposite, most like, same, and different.

(4) Gradually increase the number of options from two through five.

(5) Encourage children to examine carefully all possible responses before choosing the correct one.

(6) Progress from big pictures and words with few on a page to small pictures and words with several on a page, again until the final page is similar to an actual test page.

(7) Provide practice in putting a mark on, under, or in the circle under the correct response.

(8) Teach the children that biggest can mean the most and that a pencil may be called a marker.

(9) Use both dotted lines and heavy black lines to separate the criterion from the options.

(10) Refer to each page as a test in order to get children accustomed to the word.

(11) Encourage children to ask questions if they do not understand the test directions.

(12) Teach children to use a marker for keeping their places as they progress on each page.

(13) Gradually increase the length of time children are encouraged to remain task oriented.

The thirty pages of test wiseness materials utilized several different formats to provide children an opportunity to learn to respond to a variety of different tasks. No materials were taken from existing tests, and no attempt was made to increase the children's knowledge

in content areas measured by any readiness test. Directions given by the teachers who administered the test wiseness materials also made use of the terminology frequently found in readiness tests.

Seventy-two children (33 Negro, 36 Mexican-American, and 3 Anglo) from four Headstart classes were assigned randomly to an experimental or control group. Eleven children later were excluded due to their repeated absence or their withdrawal from the classes. Therefore, the final sample consisted of 61 children (26 Negro, 32 Mexican-American, and 3 Anglo), of which 36 were in the experimental group and 25 in the control group. The Ss in the experimental group worked with their teachers, using the test wiseness materials during thirty-minute periods twice a week for six weeks. The Ss in the control group worked with special activities directed by the teacher aides. The Metropolitan Readiness Test (MRT) was administered prior to the treatment in March (Form A), immediately after the treatment in May (Form B, post-test$_1$), and at the beginning of first grade in September (Form A, post-test$_2$).

Results

The correlation between pre- and post-test$_1$ was sufficiently high (.80) to justify using analysis of variance of the gain (or difference) scores rather than covariance analysis. Differences between pre- and post-test$_1$ were significant on Total Score (experimental> control: $F(1/60)=7.3$, $p=.01$) and on Matching (experimental > control: $F(1/60)=9.5$, $p=.003$). Differences between pre- and post-test$_2$ were not significant. Significant differences between

$post_1$-$post_2$ occurred again on Matching (control > experimental: $F(1/60)=4.7$, $p=.04$).

[put Tables 1 and 2 about here]

While the test wiseness materials were not designed to improve the test-taking abilities of a particular subgroup within the experimental sample, it was felt that the impact of the materials would be most apparent among children whose pretest scores were below the group's mean. Therefore, a mean split was made within the experimental and control groups, thus creating a low-scoring group (Ss with pre-test scores below their respective group mean). Separate analyses were made of gain scores between pre-$post_1$, pre-$post_2$, and $post_1$-$post_2$ for the two groups. Gain scores between pre-$post_1$ were greater for experimental Ss than for control Ss on Word Meaning (8.3 vs 1.0: $F(1/29)=5.0$, $p=.03$) and Matching (3.5 vs 1.1: $F(1/29)=6.6$, $p=.01$) and approached significance on Total Score (8.3 vs 2.6: $F(1/29)=3.3$, $p=.075$). All other comparisons were not significant.

Discussion

The test wiseness materials had limited value in facilitating performance on the standardized test. Significant group differences between the March and May administrations of the MRT indicate the test wiseness materials were instrumental in improving certain test-taking behaviors. The improvement was most apparent on Total Score and the Matching subtest. While other differences were not significant, mean gain scores consistently are in the expected direction.

The relative advantages initially enjoyed by the experimental group tended to dissipate over the summer months. The disproportionate gains made by the control group between May and September are not readily explainable. While the present study represents an initial attempt to identify salient prerequisite abilities pertaining to standardized tests for preschool and primary-grade children, additional work in this area is needed. Further research also is needed to examine the development of test wiseness of preschool and primary-grade children. For example, the abilities prerequisite to taking standardized tests may differ at various age and grade levels. Knowledge of these developmental patterns would facilitate test construction, evaluation, and interpretation. This knowledge also would be helpful in tailoring curricular materials to the needs of the students.

Exposure to the test wiseness materials does not necessarily ensure mastery of the objectives which underlie their development. Evidence regarding their efficacy is needed. The number of errors made on these materials provides some clues regarding their relevance. For example, many errors on the test wiseness materials would suggest that they were too difficult and that the Ss did not demonstrate mastery of the objectives. Therefore, the number of errors made on the test wiseness materials was determined and the relationship between these errors and performance on the MRT pretest was computed.

There were 208 frames on which errors could be made. The errors made were distributed throughout the frames, with some tendency for more errors to occur at the beginning of the test wiseness materials. Ss averaged a total of 33 errors over the 208 frames; that is, on the average, a S made errors on 16 percent of the frames. This figure is within the 10 to 20 percent error rate cited as being acceptable for programmed instruction materials. Therefore, some evidence exists that Ss appeared to demonstrate mastery of the materials and the objectives underlying the test wiseness instruction.

The correlation between the number of errors on the test wiseness materials and performance on the MRT pre-test was -.71. This tends to support the assumption that the test wiseness materials are related to the abilities assessed by the MRT.

No formal attempt was made to assess changes in children's attitudes regarding testing. However, a discussion was held with teachers after the administration of post-test$_1$ to discern their impressions concerning the children's attitudes. They reported that the majority of children in the experimental group appeared to be more confident on tasks requiring pencil and paper work and to remain task-oriented for longer periods of time. The use of the word "test" and being asked to take tests did not seem to disturb the children in the experimental group, while children in the control group continued to exhibit debilitating behaviors while taking tests.

Table 1

Means and Standard Deviations on Three Administrations of the Metropolitan Readiness Test

	Pre-test		Post-test		Post-test$_2$	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Experimental (N=36)						
TOTAL	38.47	12.50	44.86	13.52	48.31	12.22
Word Meaning	5.31	1.99	7.22	2.38	6.03	2.51
Listening	10.00	2.18	7.94	2.30	10.03	1.71
Matching	5.11	2.94	9.00	2.80	8.86	2.47
Alphabet	5.22	4.20	6.22	4.69	6.19	4.09
Numbers	8.33	3.18	8.72	3.50	10.03	3.17
Copying	3.94	3.54	5.69	4.17	7.17	3.75
Control (N=25)						
TOTAL	38.88	13.10	40.00	13.45	46.16	13.79
Word Meaning	5.68	1.80	6.88	2.47	5.44	1.60
Listening	10.36	2.02	7.76	2.61	9.60	2.14
Matching	5.60	3.44	6.60	2.90	8.16	3.02
Alphabet	5.44	4.30	6.96	4.45	7.04	3.66
Numbers	8.76	3.09	7.96	3.59	9.64	3.90
Copying	3.04	2.86	3.84	2.82	6.28	4.49

Table 2

Mean Gain Scores on the Metropolitan Readiness Test for Experimental and Control Groups

	Pre-Post$_1$		Pre-Post$_2$		Post$_1$-Post$_2$	
	E	C	E	C	E	C
TOTAL	6.3	1.1	4.5	6.2	9.7	7.4
Word Meaning	1.9	1.2	-1.6	-1.4	.7	-.2
Listening	-2.1	-2.6	2.1	1.8	.0	-.8
Matching	3.9	1.3	-.1	1.6	3.8	2.6
Alphabet	1.0	1.5	.0	.1	1.0	1.6
Numbers	.4	-.8	1.3	1.7	1.7	.9
Copying	1.7	.8	1.6	2.4	3.1	3.2

BIBLIOGRAPHY

Anastasi, A. Differential Psychology. New York: Macmillan, 1958.

Brazziel, William F., and Mary Terrell. An Experiment in the development of readiness in a culturally disadvantaged group of first grade children, The Journal of Negro Education, 1962, 31, 4-7.

Deutsch, M., et al., Guidelines for testing minority group children, Journal of Social Issues, 1964, 20.

Ebel, Robert L. Measuring Educational Achievement. Englewood Cliffs, New Jersey: Prentice-Hall, 1965.

Eels, Kenneth, Allison Davis, Robert Havighurst, V. E. Herrick, and Ralph Tyler. Intelligence and Cultural Differences. Chicago: University of Chicago Press, 1951.

Haggard, Ernest A. Social status and intelligence: an experimental study of certain cultural determinants of measured intelligence, Genetic Psychology Monographs, 1954, 49, 141-186.

Kreit, L. The Effects of test-taking practice on pupil test performance, Unpublished doctoral dissertation, Indiana University, 1967.

Masland, R., S. Sarason, and T. Gladwin. Mental Subnormality. New York: Basic Books, 1958.

Moore, J. C., R. E. Schutz, and R. L. Baker. The Application of a self-instructional technique to develop a test-taking strategy. American Educational Research Journal, 1966, 3, 13-17.

Moore, J. C., Manipulating the effectiveness of a self-instructional program. Journal of Educational Psychology, 1968, 59, 315-319.

Millman, J., C. Bishop, and R. Ebel. An Analysis of test wiseness, Educational and Psychological Measurement, 1965, 25, 707-727.

Millman, J., and Setijari. A Comparison of the performance of American and Indonesian students on three types of test items, Journal of Educational Research, 1966, 59, 273-275.

Newland, T. Ernest. Psychological assessment of exceptional children and youth, in William Cruickshank (ed.), Psychology of Exceptional Children and Youth. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Tyler, L. The Psychology of Individual Differences. New York: Appleton-Century-Crofts, 1956.

Slakter, Malcolm J., and Roger A. Koehler. Test Wiseness. Final Technical Report to the Teacher Education Research Center, State University College at Fredonia. 1969.

Wahlstrom, M., and F. Boersma. The Influence of test-wiseness upon achievement, Educational and Psychological Measurement, 1968, 28,